Course Syllabus

EXPLORATORY DATA ANALYSIS

Printed by: lisacabe

Program: Computer Science

1. Course number and name

CCPG1002 - EXPLORATORY DATA ANALYSIS

2. Credits and contact hours

3 credits and 3 contact hours

3. Instructor's course or coordinator's name

CARMEN KARINA VACA RUIZ

4. Text book, tittle, author, and year

• Wes McKinney. Python for Data Analysis (1st Edition)

a.Other supplemental materials

• Russell, Matthew A.. Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google+, GitHub, and More (2da Edición)

5. Specific course information

a. Brief description of the content of the course (catalog description)

This course is part of the "Big Data" Itinerary. In this course we explore subjects such as: online data,acquisition andprocessing published online to infer patterns from them. The student will be able to understand as massive data sets impact diverse science fields and our daily lives.In particular, the student will analyze and characterize data in several formats. The results of this analysis will be communicated orally and visually.

b. This course is: Selected elective

6. Specific goals for the course

a. Specific outcomes of instruction

1.- Use programming libraries for collecting data published online.

2.- Use free software to visually communicate the patterns learned from geolocated data published on online social networks.

3.- Analyze the existing correlations among several variables in the datasets in order to design predictive models.

4.- Process text corpuses using natural language processing techniques to infer knowledge regarding issues in different professions.

b. Explicity indicate which of the student outcomes listed in Criterion 3 or any other outcomes are addressed by the course

7. Brief list of topics to be covered

1.- Introduction

2.- Data Formats and sources

3.- Data Acquisition

Course Syllabus

EXPLORATORY DATA ANALYSIS

Printed by: lisacabe

Program: Computer Science

4.- Data Representation
5.- Spatial and Network data
6.- Visualization
7.- Data cleaning and transformation
8.- Data grouping and aggregation
9.- Text Processing
10.- Prediction and clustering